

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/756,283

DATE: 11/20/2001

TIME: 10:02:49

Input Set : A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

5 <110> APPLICANT: Chernajovsky, Yuti
6 Dreja, Hanna Stina
7 Adams, Gillian
10 <120> TITLE OF INVENTION: Latent Fusion Protein
13 <130> FILE REFERENCE: 0623.1000000
16 <140> CURRENT APPLICATION NUMBER: US 09/756,283
18 <141> CURRENT FILING DATE: 2001-01-09
21 <160> NUMBER OF SEQ ID NOS: 100
24 <170> SOFTWARE: PatentIn version 3.0
28 <210> SEQ ID NO: 1
30 <211> LENGTH: 15
32 <212> TYPE: PRT
C--> 34 <213> ORGANISM: Artificial
38 <220> FEATURE:
40 <223> OTHER INFORMATION: MMP cleavage site including linker sequence
42 <400> SEQUENCE: 1
44 Gly Gly Gly Gly Ser Pro Leu Gly Leu Trp Ala Gly Gly Gly Ser
45 1 5 10 15
47 <210> SEQ ID NO: 2
49 <211> LENGTH: 52
51 <212> TYPE: DNA
C--> 53 <213> ORGANISM: Artificial
57 <220> FEATURE:
59 <223> OTHER INFORMATION: Sense oligo
61 <400> SEQUENCE: 2
62 aattcggggg aggcggatcc ccgctcgggc ttgggcgagg agggggctca gc 52
65 <210> SEQ ID NO: 3
67 <211> LENGTH: 52
69 <212> TYPE: DNA
C--> 71 <213> ORGANISM: Artificial
75 <220> FEATURE:
77 <223> OTHER INFORMATION: Antisense oligo
79 <400> SEQUENCE: 3
80 ggccgctgag ccccttcccg cccaaagccc gagcggggat ccgcctcccc cg 52
83 <210> SEQ ID NO: 4
85 <211> LENGTH: 29
87 <212> TYPE: DNA
C--> 89 <213> ORGANISM: Artificial
93 <220> FEATURE:
95 <223> OTHER INFORMATION: Sense Primer
97 <400> SEQUENCE: 4
98 ccaagcttat gccgccctcc gggctgcgg 29
101 <210> SEQ ID NO: 5
103 <211> LENGTH: 29
105 <212> TYPE: DNA
C--> 107 <213> ORGANISM: Artificial
111 <220> FEATURE:

Does Not Comply
Corrected Diskette Needed

Errors on Error pp. 1 & 2

RAW SEQUENCE LISTING

DATE: 11/20/2001

PATENT APPLICATION: US/09/756,283

TIME: 10:02:49

Input Set : A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

```

113 <223> OTHER INFORMATION: Antisense Primer
115 <400> SEQUENCE: 5
116 ccgaattcgc ttgcagatg ctgggccct 29
119 <210> SEQ ID NO: 6
121 <211> LENGTH: 31
123 <212> TYPE: DNA
C--> 125 <213> ORGANISM: Artificial
129 <220> FEATURE:
131 <223> OTHER INFORMATION: Sense Primer
133 <400> SEQUENCE: 6
134 cgcggccgca atcaactata agcagctcca g 31
137 <210> SEQ ID NO: 7
139 <211> LENGTH: 32
141 <212> TYPE: DNA
C--> 143 <213> ORGANISM: Artificial
147 <220> FEATURE:
149 <223> OTHER INFORMATION: Antisense Primer
151 <400> SEQUENCE: 7
152 ggtctagatc agttttggaa gtttctggta ag 32
155 <210> SEQ ID NO: 8
157 <211> LENGTH: 29
159 <212> TYPE: DNA
C--> 161 <213> ORGANISM: Artificial
165 <220> FEATURE:
167 <223> OTHER INFORMATION: Sense Primer
169 <400> SEQUENCE: 8
170 ccaagcttat gaacaacagg tggatcctc 29
173 <210> SEQ ID NO: 9
175 <211> LENGTH: 29
177 <212> TYPE: DNA
C--> 179 <213> ORGANISM: Artificial
183 <220> FEATURE:
185 <223> OTHER INFORMATION: Antisense Primer
187 <400> SEQUENCE: 9
188 ccgaattcgt ttggaagtt tctggtaag 29
191 <210> SEQ ID NO: 10
193 <211> LENGTH: 31
195 <212> TYPE: DNA
C--> 197 <213> ORGANISM: Artificial
201 <220> FEATURE:
203 <223> OTHER INFORMATION: Sense Primer
205 <400> SEQUENCE: 10
206 cgcggccgca ctatccacct gcaagactat c 31
209 <210> SEQ ID NO: 11
211 <211> LENGTH: 32
213 <212> TYPE: DNA
C--> 215 <213> ORGANISM: Artificial
219 <220> FEATURE:
221 <223> OTHER INFORMATION: Antisense Primer

```

RAW SEQUENCE LISTING

DATE: 11/20/2001

PATENT APPLICATION: US/09/756,283

TIME: 10:02:49

Input Set : A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

```

223 <400> SEQUENCE: 11
224 ggtctagatc agctttgcag atgctgggcc ct 32
227 <210> SEQ ID NO: 12
229 <211> LENGTH: 23
231 <212> TYPE: DNA
C--> 233 <213> ORGANISM: Artificial
237 <220> FEATURE:
239 <223> OTHER INFORMATION: Sense Primer
241 <400> SEQUENCE: 12
242 cgcccatggc gccttcgggg cct 23
245 <210> SEQ ID NO: 13
247 <211> LENGTH: 29
249 <212> TYPE: DNA
C--> 251 <213> ORGANISM: Artificial
255 <220> FEATURE:
257 <223> OTHER INFORMATION: Antisense Primer
259 <400> SEQUENCE: 13
260 ccgaattcgc tgtgcaggtg ctgggccct 29
263 <210> SEQ ID NO: 14
265 <211> LENGTH: 5
267 <212> TYPE: PRT
C--> 269 <213> ORGANISM: Artificial
273 <220> FEATURE:
275 <223> OTHER INFORMATION: Flexible linker
277 <400> SEQUENCE: 14
279 Gly Gly Gly Gly Ser
280 1 5
282 <210> SEQ ID NO: 15
284 <211> LENGTH: 6
286 <212> TYPE: PRT
C--> 288 <213> ORGANISM: Artificial
292 <220> FEATURE:
294 <223> OTHER INFORMATION: Cleavage site
296 <400> SEQUENCE: 15
298 Pro Leu Gly Leu Trp Ala
299 1 5
301 <210> SEQ ID NO: 16
303 <211> LENGTH: 8
305 <212> TYPE: PRT
C--> 307 <213> ORGANISM: Artificial
311 <220> FEATURE:
313 <223> OTHER INFORMATION: Flexible portion
315 <400> SEQUENCE: 16
317 Gly Gly Gly Gly Ser Ala Ala Ala
318 1 5
320 <210> SEQ ID NO: 17
322 <211> LENGTH: 4
324 <212> TYPE: PRT
C--> 326 <213> ORGANISM: Artificial

```

RAW SEQUENCE LISTING

DATE: 11/20/2001

PATENT APPLICATION: US/09/756,283

TIME: 10:02:49

Input Set: A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

```

330 <220> FEATURE:
332 <223> OTHER INFORMATION: Core of cleavage site
334 <400> SEQUENCE: 17
336 Pro Leu Gly Leu
337 1
339 <210> SEQ ID NO: 18
341 <211> LENGTH: 4
343 <212> TYPE: PRT
C--> 345 <213> ORGANISM: Artificial
349 <220> FEATURE:
351 <223> OTHER INFORMATION: Core of cleavage site
353 <400> SEQUENCE: 18
355 Pro Leu Gly Ile
356 1
358 <210> SEQ ID NO: 19
360 <211> LENGTH: 1376
362 <212> TYPE: DNA
C--> 364 <213> ORGANISM: Artificial
368 <220> FEATURE:
370 <223> OTHER INFORMATION: LAP-mIFNbeta construct
372 <220> FEATURE:
374 <221> NAME/KEY: CDS
376 <222> LOCATION: (1)..(1368)
380 <400> SEQUENCE: 19
381 atg ccg ccc tcc ggg ctg cgg ctg ctg ccg ctg ctg cta ccg ctg ctg      48
382 Met Pro Pro Ser Gly Leu Arg Leu Leu Pro Leu Leu Leu Pro Leu Leu
383 1                    5                10                15
385 tgg cta ctg gtg ctg acg cct ggc ccg ccg gcc gcg gga cta tcc acc      96
386 Trp Leu Leu Val Leu Thr Pro Gly Pro Pro Ala Ala Gly Leu Ser Thr
387                20                25                30
389 tgc aag act atc gac atg gag ctg gtg aag cgg aag cgc atc gag gcc      144
390 Cys Lys Thr Ile Asp Met Glu Leu Val Lys Arg Lys Arg Ile Glu Ala
391                35                40                45
393 atc cgc ggc cag atc ctg tcc aag ctg cgg ctc gcc agc ccc ccg agc      192
394 Ile Arg Gly Gln Ile Leu Ser Lys Leu Arg Leu Ala Ser Pro Pro Ser
395                50                55                60
397 cag ggg gag gtg ccg ccc ggc ccg ctg ccc gag gcc gtg ctc gcc ctg      240
398 Gln Gly Glu Val Pro Pro Gly Pro Leu Pro Glu Ala Val Leu Ala Leu
399 65                70                75                80
401 tac aac agc acc cgc gac cgg gtg gcc ggg gag agt gca gaa ccg gag      288
402 Tyr Asn Ser Thr Arg Asp Arg Val Ala Gly Glu Ser Ala Glu Pro Glu
403                85                90                95
405 ccc gag cct gag gcc gac tac tac gcc aag gag gtc acc cgc gtg cta      336
406 Pro Glu Pro Glu Ala Asp Tyr Tyr Ala Lys Glu Val Thr Arg Val Leu
407                100               105               110
409 atg gtg gaa acc cac aac gaa atc tat gac aag ttc aag cag agt aca      384
410 Met Val Glu Thr His Asn Glu Ile Tyr Asp Lys Phe Lys Gln Ser Thr
411                115               120               125
413 cac agc ata tat atg ttc ttc aac aca tca gag ctc cga gaa gcg gta      432

```

RAW SEQUENCE LISTING

DATE: 11/20/2001

PATENT APPLICATION: US/09/756,283

TIME: 10:02:49

Input Set : A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

```

414 His Ser Ile Tyr Met Phe Phe Asn Thr Ser Glu Leu Arg Glu Ala Val
415      130      135      140
417 cct gaa ccc gtg ttg ctc tcc cgg gca gag ctg cgt ctg ctg agg agg      480
418 Pro Glu Pro Val Leu Leu Ser Arg Ala Glu Leu Arg Leu Leu Arg Arg
419 145      150      155      160
421 ctc aag tta aaa gtg gag cag cac gtg gag ctg tac cag aaa tac agc      528
422 Leu Lys Leu Lys Val Glu Gln His Val Glu Leu Tyr Gln Lys Tyr Ser
423      165      170      175
425 aac aat tcc tgg cga tac ctc agc aac cgg ctg ctg gca ccc agc gac      576
426 Asn Asn Ser Trp Arg Tyr Leu Ser Asn Arg Leu Leu Ala Pro Ser Asp
427      180      185      190
429 tcg cca gag tgg tta tct ttt gat gtc acc gga gtt gtg cgg cag tgg      624
430 Ser Pro Glu Trp Leu Ser Phe Asp Val Thr Gly Val Val Arg Gln Trp
431      195      200      205
433 ttg agc cgt gga ggg gaa att gag ggc ttt cgc ctt agc gcc cac tgc      672
434 Leu Ser Arg Gly Gly Glu Ile Glu Gly Phe Arg Leu Ser Ala His Cys
435      210      215      220
437 tcc tgt gac agc agg gat aac aca ctg caa gtg gac atc aac ggg ttc      720
438 Ser Cys Asp Ser Arg Asp Asn Thr Leu Gln Val Asp Ile Asn Gly Phe
439 225      230      235      240
441 act acc ggc cgc cga ggt gac ctg gcc acc att cat ggc atg aac cgg      768
442 Thr Thr Gly Arg Arg Gly Asp Leu Ala Thr Ile His Gly Met Asn Arg
443      245      250      255
445 cct ttc ctg ctt ctc atg gcc acc ccg ctg gag agg gcc cag cat ctg      816
446 Pro Phe Leu Leu Leu Met Ala Thr Pro Leu Glu Arg Ala Gln His Leu
447      260      265      270
449 caa agc gaa ttc ggg gga ggc gga tcc ccg ctc ggg ctt tgg gcg gga      864
450 Gln Ser Glu Phe Gly Gly Gly Gly Ser Pro Leu Gly Leu Trp Ala Gly
451      275      280      285
453 ggg ggc tca gcg gcc gca atc aac tat aag cag ctc cag ctc caa gaa      912
454 Gly Gly Ser Ala Ala Ala Ile Asn Tyr Lys Gln Leu Gln Leu Gln Glu
455      290      295      300
457 agg acg aac att cgg aaa tgt cag gag ctc ctg gag cag ctg aat gga      960
458 Arg Thr Asn Ile Arg Lys Cys Gln Glu Leu Leu Glu Gln Leu Asn Gly
459 305      310      315      320
461 aag atc aac ctc acc tac agg gcg gac ttc aag atc cct atg gag atg      1008
462 Lys Ile Asn Leu Thr Tyr Arg Ala Asp Phe Lys Ile Pro Met Glu Met
463      325      330      335
465 acg gag aag atg cag aag agt tac act gcc ttt gcc atc caa gag atg      1056
466 Thr Glu Lys Met Gln Lys Ser Tyr Thr Ala Phe Ala Ile Gln Glu Met
467      340      345      350
469 ctc cag aat gtc ttt ctt gtc ttc aga aac aat ttc tcc agc act ggg      1104
470 Leu Gln Asn Val Phe Leu Val Phe Arg Asn Asn Phe Ser Ser Thr Gly
471      355      360      365
473 tgg aat gag act att gtt gta cgt ctc ctg gat gaa ctc cac cag cag      1152
474 Trp Asn Glu Thr Ile Val Val Arg Leu Leu Asp Glu Leu His Gln Gln
475      370      375      380
477 aca gtg ttt ctg aag aca gta cta gag gaa aag caa gag gaa aga ttg      1200
478 Thr Val Phe Leu Lys Thr Val Leu Glu Glu Lys Gln Glu Glu Arg Leu

```

VERIFICATION SUMMARY

DATE: 11/20/2001

PATENT APPLICATION: US/09/756,283

TIME: 10:02:50

Input Set : A:\0623 1000000 seq list.txt

Output Set: N:\CRF3\11202001\I756283.raw

L:34 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1
L:53 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2
L:71 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:89 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:107 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:125 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:143 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:161 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:179 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:197 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:215 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:233 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:251 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:269 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:288 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:307 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:326 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:345 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18
L:364 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19
L:504 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:508 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:508 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:632 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:770 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:774 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:774 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:1604 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1659 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1759 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55

<210> 20

<211> 455

<212> PRT

<213> Artificial[↑]

→ Must provide explanation
— See error summary sheet, item 11

<400> 20

Met	Pro	Pro	Ser	Gly	Leu	Arg	Leu	Leu	Pro	Leu	Leu	Leu	Pro	Leu	Leu
1				5					10					15	

04/756483

error 1.2

<210> 22

<211> 447

<212> PRT

<213> Artificial

→ same error

<400> 22

Met	Asn	Asn	Arg	Trp	Ile	Leu	His	Ala	Ala	Phe	Leu	Leu	Cys	Phe	Ser
1				5					10					15	

Raw Sequence Listing Error Summary

ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/756,283

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics The number/text at the end of each line "wrapped" down to the next line. This may occur if your file
 Wrapped Aminos was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will
 prevent "wrapping."

- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.

- 3 Misaligned Amino The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers;
 Numbering use space characters, instead.

- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please
 ensure your subsequent submission is saved in ASCII text.

- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules,
 each n or Xaa can only represent a single residue. Please present the maximum number of each
 residue having variable length and indicate in the <220>-<223> section that some may be missing.

- 6 PatentIn 2.0 A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
 "bug" sequences(s) . Normally, PatentIn would automatically generate this section from the
 previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to
 the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for
 Artificial or Unknown sequences.

- 7 Skipped Sequences Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (OLD RULES) (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

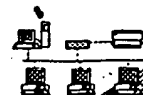
- 8 Skipped Sequences Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 (NEW RULES) <210> sequence id number
 <400> sequence id number
 000

- 9 Use of n's or Xaa's Use of n's and/or Xaa's have been detected in the Sequence Listing.
 (NEW RULES) Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

- 10 Invalid <213> Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or
 Response scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or
 is Artificial Sequence

- 11 ✓ Use of <220> Sequence(s) 20+22 missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or
 "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

- 12 PatentIn 2.0 Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file,
 "bug" resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence
 listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.



RAW SEQUENCE LISTING **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/756,283

Source: OPE

Date Processed by STIC: 11/20/01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>